

REMARKS

Applicants respectfully request reconsideration and allowance in view of the foregoing amendments and following remarks. In the Office Action, mailed December 22, 2003, the Examiner rejected claims 1-39. By this amendment, claim 17 has been cancelled. Following entry of these amendments, claims 1-16 and 18-39 will be pending in the application.

Claim Rejections under 35 U.S.C. § 112 (Second Paragraph)

In the Office Action, the Examiner rejected claims 2, 17, 20 and 26 under 35 U.S.C. §112, second paragraph, for allegedly being indefinite for failing to particularly point out and distinctly claim the subject matter that Applicants regard as the invention. Specifically, the Examiner noted as indefinite the phrases "determine if there are enough documents," "enough documents" and "where there is too much overlap." Applicants respectfully traverse these rejections.

A rejection under 35 U.S.C. §112, second paragraph, is not appropriate when "those skilled in the art would understand what is claimed when the claim is read in light of the specification." *Orthokinetics, Inc. v. Safety Travel Chairs, Inc.*, 806 F.2d 1565, 1576 (Fed. Cir. 1986). In the Office Action, the Examiner asserts that claims 2, 17, 20 and 26 are formulated such that it is unclear what Applicants mean by "determine if there are enough documents," "enough documents" and "where there is too much overlap." Applicants respectfully submit that the specification fully defines and supports these phrases. Moreover, those skilled in the art would understand these phrases in the context of the present invention after being taught by the specification.

Regarding the "enough documents" type phrases of claims 2, 20 and 26, Applicants' specification as filed at page 12, last three paragraphs, discloses the details of how one embodiment makes the decision of whether there are enough or not enough documents in the temporary category to either construct a new category for the documents or assign them to the non-specific category. Thus, those skilled in the art would readily understand what is claimed in light of Applicants' specification.

Regarding the "too much overlap" type phrase of claim 17, Applicants' specification as filed fully supports this phraseology and those skilled in the art would therefore completely understand what is claimed in light of Applicants' disclosure. Applicant has cancelled claim 17 rendering the

Examiner's rejection moot.

For at least the foregoing reasons, the §112, second paragraph, rejections of claims 2, 17, 20 and 26 should be withdrawn.

Claim Rejections under 35 U.S.C. §103(a)

In the Office Action, the Examiner rejected claims 1-16 and 18-39 under 35 U.S.C. §103(a) as allegedly being unpatentable over U.S. Patent No. 6,237,011 to Ferguson, et al. (hereinafter "Ferguson") in view of U.S. Patent No. 6,571,240 to Ho, et al. (hereinafter "Ho"). Further, the Examiner rejected claim 17 under 35 U.S.C. §103(a) as allegedly being unpatentable over the combination of Ferguson and Ho, and further in view of the article entitled "Grouper: A Dynamic Clustering Interface to Web Search Results" by Zamir, et al. (hereinafter "Zamir"). Applicants respectfully traverse the rejections of claims 1-39 and note the following standards for a proper §103(a) rejection.

A §103(a), or obviousness, rejection is proper only when "the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which the subject matter pertains." 35 U.S.C. §103(a). The Examiner must make out a *prima facie* case for obviousness. The *en banc* Federal Circuit has held that "structural similarity between claimed and prior art subject matter, proved by combining references or otherwise, where the prior art gives reason or motivation to make the claimed compositions, creates a *prima facie* case of obviousness." *In re Dillon*, 16 U.S.P.Q. 2d 1897, 1901 (CAFC 1990).

Additionally, the mere fact that references can be combined or modified does not render the resultant combination obvious unless the prior art also suggests the desirability of the combination. *In re Mills*, 916 F.2d 680, 16 U.S.P.Q.2d 1430 (Fed. Cir. 1990). Likewise, if the proposed modification would render the prior art invention being modified unsatisfactory for its intended purpose, then there is no suggestion or motivation to make the proposed modification. *In re Gordon*, 733 F.2d 900, 221 U.S.P.Q. 1125 (Fed. Cir. 1984).

Further, with hindsight, a claim of obviousness can be an easy one to make. Many inventions seem obvious with the clarity of 20-20 hindsight. However, a hindsight basis for

obviousness is inappropriate and cannot sustain a *prima facie* case of obviousness.

For at least the reasons stated below and taking into consideration the standards for obviousness presented above, Applicants assert that one of ordinary skill in the art would not have considered Applicants' invention obvious at the time of invention and, therefore, that Applicants' rejected claims 1-39 are not obvious over the prior art of record.

Claim 1

Applicants' independent claim 1 recites a method of categorizing an initial collection of documents, each document being represented by a string of characters, that includes the steps of:

- identifying predefined characters in the string of characters from the documents in the initial collection of documents to form identified characters;
- changing the identified characters in the documents in the initial collection of documents to form a preprocessed collection of documents, each of the preprocessed collection of documents represented by a preprocessed string of characters;
- constructing a number of categories from the preprocessed string of characters of the preprocessed collection of documents; and
- assigning each document in the preprocessed collection of documents to a category to form a hierarchy of categories of documents.

In rejecting Applicants' independent claim 1, the Examiner refers to col. 8, ll. 12-32 of Ferguson, and col. 12, ll. 16-26 of Ho. The document categorization method and apparatus disclosed by Ferguson uses pre-defined data storage files ("STG files") that contain pre-populated criteria (e.g., a copy of the text of the document) that has previously been indexed to match documents by pre-indexed criteria (Ferguson, col. 3, ll. 60-65 and col. 8, ll. 12-32). Ho describes a method of regularizing words in a document only by replacing them with their roots (Ho, col. 12, l. 19). Nowhere does Ferguson or Ho disclose or suggest assigning documents to categories constructed from the preprocessed string of characters that represent each of a preprocessed collection of documents formed from the documents as required by independent claim 1.

In contrast, claim 1 of the present invention discloses a categorization method that does not require the pre-indexing, as does Ferguson, and preprocesses the documents in a manner dissimilar to that disclosed in Ho. Claim 1 of the present invention categorizes the documents based on the preprocessed string of characters that make up the documents themselves. Applicants' process

categorizes one document at a time, to create one category at a time, without first creating a category structure for the documents. Claim 1 of the present invention performs the preprocessing of the documents by removing stop-characters and stop-words, and converting upper case characters and abbreviations, as well as converting words to their root form (Applicants' Specification, pp. 7-9 in relation to Figure 3).

Specification, the Ferguson method first creates a data storage file ("STG file") for every document in its database. The STG file contains standardized fields and will be created, modified and deleted as the corresponding documents in the database are created, modified and deleted. See, Ferguson, col. 3, ll. 60-65. One of the standardized fields of the Ferguson STG file "reflects the raw text data associated with the corresponding document," where the "raw text data is primarily used for indexing purposes." See, Ferguson, col. 4, ll. 13-15. Further, when an STG file is created (or modified), the Ferguson index engine "creates a new entry [modifies an existing entry] in the indexing database for the new [modified] document, cross-referenced with key terms and other attributes extracted from the new [modified] document's STG file." See, Ferguson, col. 6, ll. 22-25.

In addition to using the STG file for pre-indexing, Ferguson uses the STG file along with a search function. The search function of Ferguson "employs a search engine that globally searches the document collection (i.e., the STG files in the STG file directory) and retrieves documents that fit or match a number of user-defined conditions with respect to text, meta-text, and/or other file attributes." See, Ferguson, col. 12, ll. 59-64. Thus, for a user to be able to find information desired and/or believed to exist within the database, the STG files must maintain an accurate textual duplication of the documents. That is one reason why modifications to documents within the Ferguson system are reflected in the corresponding STG files. See, Ferguson, col. 4, ll. 50-54. Thus, there is no motivation at all for someone skilled in the art to combine Ho to Ferguson. In fact, the combination of the two would render the Ferguson search function incapable of performing in its intended fashion.

Specifically, the Ho method discloses a hashing function to convert words into their root forms. The Ho method can be used to convert plural nouns to singular and verbs of varying forms to the present tense. See, Ho, col. 12, ll. 60-64. In contrast, the preprocessing of the documents in Applicants' invention includes one or more of the following: removing stop-characters (e.g.,

punctuation marks, foreign character, etc.), removing stop-words (e.g., definite articles, indefinite articles, etc.), converting upper case characters to lower case and converting abbreviations to full-length words. See, Applicants' Specification, pp. 7-9 in relation to Figure 3. Applicants' preprocessing might further include converting words to their root form. Thus, Applicants' preprocessing content is patentably distinct from Ho. Additionally, Applicants' preprocessing methodology is patentably distinct from Ho because Applicants' invention does not use a hashing function to convert words to root form. Rather, Applicants' invention uses one or more dictionary-type look-up or comparison files stored within the system to preprocess the documents.

Therefore, for at least the reasons presented above, Applicants request the withdrawal and reconsideration of the claim rejections for independent claim 1. Applicants respectfully submit that independent claim 1 is in a condition for allowance, and respectfully request such a Notice to that effect.

Dependent Claims 2-25, 33 and 35

Dependent claims 2-16, 18-25, 33 and 35 [claim 17 having been cancelled] all ultimately depend from independent claim 1. The allowability of dependent claims 2-16, 18-25, 33 and 35 thus follows from the allowability of independent claim 1; as such, dependent claims 2-16, 18-25, 33 and 35 are allowable over the art of record.

Specifically, regarding dependent claim 2, the Examiner refers to col. 8, ll. 12-32 of Ferguson to support the rejection. Nowhere does Ferguson suggest or disclose the testing step, the constructing step or the assigning step of Applicants' claimed invention. Ferguson does not test a collection of similar documents in a temporary grouping to determine whether a new category should be constructed. Rather, Ferguson assigns a folder with the attributes of a document and then places similar documents in that folder, if there are any. Thus, since Ferguson creates folders regardless of the number of documents that might be assigned to that folder, Ferguson does not construct a category if the quantity of related documents dictates, nor does Ferguson assign related documents to a temporary category if their number is minimal.

Specifically, regarding dependent claim 3-7, the Examiner has made a leap of faith that because Ho discloses a hashing function to change words to their root form, one of ordinary skill in the art would have known to make the changes in Applicants' claims 3-7. Applicants assert,

though, that user readability is reduced by converting capitals to lower case letters, removing in/definite articles, etc., especially after punctuation marks have been removed. Making these changes, as already discussed above, defeats one of the express aspects of Ferguson, the search function.

Specifically, regarding dependent claim 17, Applicants' have cancelled claim 17 rendering the Examiner's rejection moot.

Specifically, regarding dependent claim 20, neither Ho nor Ferguson teaches or suggests the use of a temporary category as claimed by Applicants. Applicants' temporary category, as claimed, does not have category properties assigned. It is merely a holding area, or testing area. All categories in Ferguson are determined by, and have the properties of, the initial STG file of the document placed within it. Thus, neither Ho nor Ferguson contemplate the idea of a temporary category.

Specifically, regarding dependent claims 23-25, neither Ferguson nor Ho teaches or suggests the use of an anchor-text character string as in Applicants' claimed invention.

Therefore, for at least these reasons presented above, Applicants respectfully submit that dependent claims 2-16, 18-25, 33 and 35 are in a condition for allowance, and respectfully request such a Notice to that effect.

Claims 26-32, 34 and 36-39

In the office action, the Examiner rejected claims 26-39 "on grounds corresponding to the reasons given above for claims 1-16 and 18-25." Therefore, Applicants contend that, for at least all of the reasons for allowability presented above in relation to the rejections of claims 1-25, 33 and 35, the art of record neither discloses nor suggests the subject matter of claims 26-32, 34 and 36-39; thus, these claims are allowable over the art of record.

Therefore, for at least these reasons, Applicants respectfully submit that claims 26-32, 34 and 36-are in a condition for allowance, and respectfully request such a Notice to that effect.


Conclusion

All objections and rejections having been addressed, it is respectfully submitted that the

present application is in a condition of allowance and a Notice to that effect is earnestly solicited. If any points remain in issue which the Examiner feels may be best resolved through a personal or telephone interview, the Examiner is kindly requested to contact the undersigned at the telephone number listed below.

CHARGE STATEMENT: The Commissioner is hereby authorized to charge fees that may be required relative to this application, or credit any overpayment, to our Account 03-3975, Order No. 053684-0300105 (LS-002).

Respectfully submitted,
PILLSBURY WINTHROP LLP

By: 
Ross L. Franks, Reg. No. 47,233
For: David A. Jakopin, Reg. No. 32,995

2550 Hanover Street
Palo Alto, CA 94304-1115
Tel. No.: (650) 233-4790
Fax No.: (650) 233-4545

60366630